



## An optimization approach for aerosol retrievals using simulated MISR radiances

David J. Diner <sup>a,\*</sup>, Rachel A. Hodos <sup>a</sup>, Anthony B. Davis <sup>a</sup>, Michael J. Garay <sup>b</sup>, John V. Martonchik <sup>a</sup>, Suniti V. Sanghavi <sup>a</sup>, Paul von Allmen <sup>a</sup>, Alexander A. Kokhanovsky <sup>c</sup>, Pengwang Zhai <sup>d</sup>

<sup>a</sup> Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, United States

<sup>b</sup> Raytheon Company, Pasadena, CA, United States

<sup>c</sup> University of Bremen, Bremen, Germany

<sup>d</sup> Science Systems and Applications, Inc., Hampton, VA, United States

### ARTICLE INFO

#### Article history:

Received 14 December 2010

Received in revised form 6 May 2011

Accepted 27 May 2011

#### Keywords:

Aerosols

Optimization

MISR

### ABSTRACT

Currently, many satellite-based aerosol retrievals make use of lookup tables (LUTs) containing precomputed solutions to the radiative transfer (RT) equation. The benefit of this strategy is the avoidance of expensive runtime calculations, but its main drawback is that the LUTs discretize what is inherently a continuous, multivariate solution space. The operational retrieval algorithm for the Multi-angle Imaging SpectroRadiometer (MISR), for example, compares the observations to a set of 74 aerosol mixtures, each composed of particle models having prescribed optical properties and size distributions. In a recent “blind” study comparing the performance of several satellite retrieval algorithms on simulated data over a black surface, the MISR algorithm performed reasonably well in recovering the “true” spectral aerosol optical depths (AODs), but because the correct aerosol model was not contained within the MISR LUT, the retrieved AODs were biased low by ~14%. This motivated an investigation of whether an optimization approach, in which the aerosols are modeled by a set of continuously variable parameters recovered using nonlinear least-squares, could improve the results. In this paper, we demonstrate that such an approach using Levenberg–Marquardt optimization yields superior accuracy. Advances in computer speed, development of more efficient RT codes, and algorithm innovations will be necessary for this approach to satisfy the demands of a global, production-level satellite aerosol retrieval process, especially when used in conjunction with future instruments having enhanced sensitivity to diverse aerosol properties.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

The regional and global perspectives afforded by aircraft and satellites establish a key role for remote sensing in measuring the distribution, abundance, and microphysical properties of tropospheric aerosols. Retrievals from Earth orbit using measurements of backscattered sunlight have progressed from observations employing a single spectral

band and view direction to more comprehensive methods employing combinations of multispectral, multiangular, and/or polarimetric observations (see Lee et al., 2009 for a review). The rationale behind this increasing level of measurement sophistication is the fact that the retrieval problem is inherently underdetermined, with confounding factors such as particle shape, size distribution, complex refractive index, vertical distribution, and surface reflection affecting the determination of the column-integrated aerosol optical depth (AOD).

Sensitivities to the characteristics of the aerosol optical and microphysical models used in various satellite retrieval algorithms vary from one sensor to another, as do the cloud

\* Corresponding author at: JPL MS 169-237, 4800 Oak Grove Drive, Pasadena, CA 91109 United States.

E-mail address: [David.J.Diner@jpl.nasa.gov](mailto:David.J.Diner@jpl.nasa.gov) (D.J. Diner).

screening methodologies, instrument calibration accuracies, and sensitivities to the reflectance of the underlying surface (e.g., Kokhanovsky et al., 2007; Li et al., 2009). To isolate the effect of aerosol model assumptions from the other factors, Kokhanovsky et al. (2010a) (henceforth K2010), undertook a “blind” retrieval study in which synthetic radiances simulating measurements from a variety of satellite sensors were provided to several algorithm teams. Teams representing the following instruments participated: MODIS (MODerate resolution Imaging Spectroradiometer), MERIS (MEdium Resolution Imaging Spectrometer), AATSR (Advanced Along-Track Scanning Radiometer), MISR (Multi-angle Imaging SpectroRadiometer), and POLDER (POLarization and Directionality of Earth's Reflectances). The simulated observations corresponded to an idealized situation: black surface, no clouds, no measurement noise, and observations in the principal scattering plane. At the outset, none of the retrieval teams knew the properties of the aerosol model used to generate the simulated measurements, nor did they know the optical depths used in each of the 16 model cases. Differences among the results obtained by the various teams, and between the results and the input “truth”, were intended to provide a measure of robustness of the retrieval algorithms and their associated aerosol model assumptions.

Among the key conclusions of the study were the following: (1) single-angle observations, even those with broad spectral coverage in the 400–900 nm range, did not sufficiently constrain the aerosol type for the K2010 aerosol model with the consequence that accurate AOD results were obtained only if the correct model was prescribed, (2) employing a multiangular observing strategy with intensity-only observations improved the accuracy of the results significantly due to increased sensitivity to aerosol type, and (3) the best results were obtained with a multiangular spectropolarimetric strategy. In this paper, we revisit conclusions (2) and (3) in light of the retrieval results obtained from the simulated MISR observations. The standard MISR retrieval algorithm compares the observations to a set of 74 aerosol mixtures, each composed of particle models having prescribed optical properties and size distributions (Kahn et al., 2010). For computational efficiency, the top-of-atmosphere (TOA) radiances corresponding to the component aerosol models are stored in a lookup table (LUT). In the course of the original study, we found that none of the standard set of 74 aerosol models met the goodness-of-fit criteria for a successful retrieval. Better fits were obtained by expanding the solution space of the LUT, though still using mixtures with discrete values of particle size and prescribed real refractive index. The resulting AODs, in the range 0.0–5.0, matched the input “truth” to within  $\pm 0.05 \pm 0.15\tau$  (where  $\tau$  denotes AOD) and the variation of AOD with wavelength was reproduced, though the retrieved AODs were systematically biased low by ~14%. Systematic differences between the simulated observations and the radiances predicted by the best-fitting model strongly suggested that an optimal solution had not yet been obtained. Given the inherent limitations and discretizations implicit in the use of LUTs (Katsev et al., 2009), these results led us to pose the following question: could the MISR retrievals be improved if a more robust means of exploring the solution space was employed, or was the residual bias due to an inherent lack of information content in the simulated radiances?

In this paper, we answer this question by replacing the LUT with an approach based on least-squares optimization, and demonstrate that this method leads to superior retrieval results. Following a review in Section 2 of the simulated MISR retrievals from the K2010 study, Section 3 describes the particle scattering and radiative transfer (RT) calculations used in the optimization algorithm. The algorithm itself is described in Section 4 and the corresponding aerosol retrievals are discussed in Section 5. Conclusions are presented in Section 6.

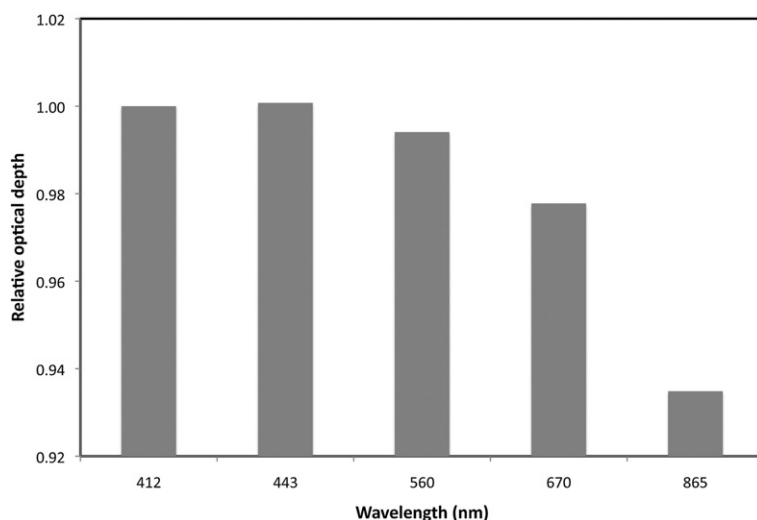
## 2. Review of previous study

The aerosol model used to generate the synthetic TOA observations assumed nonabsorbing particles (single scattering albedo = 1.0) with real refractive index of 1.38 at 400–900 nm. Letting  $r$  denote particle radius and  $N$  the total number of particles per unit volume, the particle size distribution was modeled as lognormal:

$$n(r) = \frac{N}{\sqrt{2\pi r \ln \sigma}} \exp \left[ \frac{-\ln^2(r/r_m)}{2 \ln^2 \sigma} \right]. \quad (1)$$

The median radius of the aerosol particles,  $r_m$ , was set to 0.1  $\mu\text{m}$  and  $\ln \sigma$  (denoted by the variable  $s$  in K2010) was set to unity (i.e.,  $\sigma = e = 2.718$ ). The distribution was then truncated to minimum and maximum radii  $r_{\min} = 0.05 \mu\text{m}$ ,  $r_{\max} = 20 \mu\text{m}$ , resulting in effective radius  $r_{\text{eff}}$  and effective size variance  $v_{\text{eff}}$  for this model of 1.2  $\mu\text{m}$  and 1.5, respectively. Forward RT calculations were performed using the software package SCIATRAN (Rozanov et al., 2005). The vertical profile of molecular scattering was taken from the SCIATRAN database (for 45°N latitude). The Rayleigh optical depth was 0.2376, 0.09097, 0.04391 and 0.01564 for MISR's blue, green, red and near-IR channels, respectively, and the Rayleigh depolarization factor was set to 0.0295, independent of wavelength. A homogeneous plane-parallel aerosol layer located in the altitude range 0–2 km was assumed and the Rayleigh-aerosol atmosphere was stratified into sublayers 1 km thick. The solar zenith angle was fixed at 60° and the surface was identified as being completely black.

Synthetic TOA bidirectional reflectance factors (BRFs) were supplied to the JPL MISR team (and other teams) for 16 test cases. BRF is given by  $\pi I / \mu_0 F_0$ , where  $I$  is the intensity,  $\mu_0$  is the cosine of the solar zenith angle, and  $F_0$  is the exo-atmospheric solar irradiance. For MISR, BRFs were supplied at 9 nominal viewing angles (0.0°, 26.1°, 46.1°, 60.0°, and 70.5°) in the principal plane (0° and 180° relative azimuth), in four wavelengths—443 nm (blue), 560 nm (green), 670 nm (red), and 865 nm (near-IR). Spectral AODs of the K2010 aerosol model at these wavelengths, relative to the AOD at the reference wavelength of 412 nm ( $\tau_{412}$ ) are shown in Fig. 1. The band centers are close to the values for the actual MISR instrument (446, 558, 672, and 867 nm) (Diner et al., 1998). The 16 cases for which simulated BRFs were provided ranged from a pure molecular scattering atmosphere with no aerosol (Case 1); five cases in which  $\tau_{412}$  ranged from 0.01 to 0.05 in steps of 0.01 (Cases 2–6); five cases in which  $\tau_{412}$  ranged from 0.1 to 0.5 in steps of 0.1 (Cases 7–11); and five cases in which  $\tau_{412}$  ranged from 1.0 to 5.0 in steps of 1.0 (Cases 12–16).



**Fig. 1.** Optical depth of the aerosol model used in K2010 relative to the value at 412 nm, for the four spectral bands used to simulate MISR observations.

These AOD values, their spectral dependence, and the corresponding aerosol particle parameters were not revealed to the retrieval teams at the outset of the study.

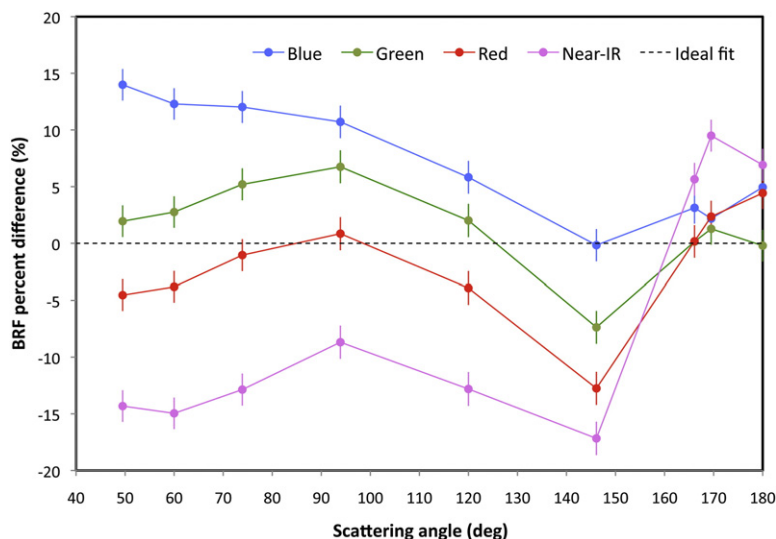
Analysis of the simulated MISR BRF dataset was performed using a modified version of the MISR Research Aerosol Retrieval algorithm described by Kahn et al. (2001, 2007). Like the MISR operational retrievals (Martonchik et al., 1998), the code is based on the Grant and Hunt (1968) Matrix Operator approach combined with modified linear mixing (Abdou et al., 1997) and integrated with a Mie code similar to that of Mishchenko et al. (2002). This RT code is scalar but includes a polarization correction for Rayleigh scattering. The Research Aerosol Retrieval was used in this investigation because the operational MISR code is not designed to handle specialized datasets such as the one supplied for the K2010 study.

Operationally, MISR aerosol retrievals are performed by comparing observed equivalent reflectances (BRF times  $\mu_0$ ) to TOA equivalent reflectances stored in the LUT. These reflectances are generated from forward RT calculations using prespecified aerosol models. The current version of the MISR aerosol retrieval (Version 22) uses a set of 74 aerosol mixtures that are made up of combinations of up to three individual aerosol components. These “pure particle” components represent a single aerosol type (e.g., sulfate) with a specified refractive index and lognormal size distribution. The 74 aerosol mixtures are based on a total of eight pure particle components: three non-absorbing spherical aerosol models with particle effective radii of 0.06, 0.26, and 2.80  $\mu\text{m}$ ; three spherical aerosol models with particle effective radius of 0.12  $\mu\text{m}$  and green band single scattering albedos of 1.0, 0.9, and 0.8; and two nonspherical dust models. The height-dependent partitioning between Rayleigh and aerosol scattering by spherical particles is represented by two exponential distributions (8 km scale height for molecules, 2 km for aerosols). Dust particles are assumed to be uniformly distributed in a lofted layer between 3 and 6 km. The Research Aerosol Retrieval was configured to take each of the 74 aerosol mixtures and generate a LUT based on the

given geometry with each mixture incremented in green band aerosol optical depth (AOD) in steps of 0.02 in the range 0.0 to 5.0. For each mixture, the AOD that minimizes a goodness of fit metric is sought. Then, a parabolic fit of the AOD to the goodness of fit metric is performed to generate a final estimate of AOD. Additional goodness of fit metrics are then calculated based on this retrieved AOD, as described in Martonchik et al. (1998).

Of the 74 input aerosol mixtures, mixture 26 was the most frequently passing mixture and was the best fitting mixture in 10 of the 16 simulation cases. This mixture consists of two components. The first accounts for 60% of the total aerosol optical depth in the green band, and has an effective particle radius of 0.26  $\mu\text{m}$ , a  $\sigma$  (antilogarithm of the lognormal size distribution standard deviation) of 1.75, and a spectrally neutral refractive index of 1.45–0.0i (i.e., non-absorbing). The second component accounts for the other 40% of the total aerosol optical depth, and has an effective particle radius of 2.80  $\mu\text{m}$ , a  $\sigma$  of 1.90, and the same refractive index. The size parameters for the mixture as a whole are  $r_{\text{eff}}=1.25 \mu\text{m}$ ,  $v_{\text{eff}}=1.99$ . Despite mixture 26 being the “best-fitting” model from the operational MISR set, the fit between the BRFs calculated from mixture 26 and the simulated observations is actually quite poor; an example is shown in Fig. 2 for Case 12 (for which  $\tau_{412}=1.0$ ).

As shown in Fig. 2, the percent differences between the provided and modeled BRFs are quite large compared to the uncertainties in the MISR measurements, which are indicated by the error bars. For Case 12 the BRFs range between 0.082 and 1.010, depending on angle and band. At these brightness levels, the MISR signal-to-noise ratio (SNR) contributes a measurement uncertainty of <0.2% at the spatial resolution used for MISR aerosol retrievals over water; additional sources of instrument relative calibration errors raise the uncertainty to 1.5% (Bruegge et al., 1998). These prelaunch performance expectations have been validated by in-flight analyses (Bruegge et al., 2002). In Fig. 2 and subsequent figures of similar content, the error bars for a given BRF are calculated from the tabulated data in Bruegge et al. (1998) for a



**Fig. 2.** Percent difference in BRFs from MISR standard mixture 26 and the simulated observations for K2010 Case 12, plotted versus scattering angle. The error bars are calculated from the tabulated data in Bruegge et al. (1998) for a measurement spatial resolution of  $1.1 \text{ km} \times 1.1 \text{ km}$ .

measurement spatial resolution of  $1.1 \text{ km} \times 1.1 \text{ km}$  (corresponding to a  $4 \times 4$  average of data from individual  $275 \text{ m} \times 275 \text{ m}$  pixels), which is the spatial averaging mode used in MISR aerosol retrievals over water. Absolute calibration biases, estimated at 3–4% for MISR (Bruegge et al., 2002) are not included in the estimated uncertainties; such biases can affect the results from any retrieval scheme (LUT, optimization, or other) and we consider only relative channel-to-channel errors and the contribution of instrument noise to be the relevant uncertainties for comparing the optimization and LUT-based results. The large BRF differences shown in Fig. 2 imply that mixture 26 would not have met the standard MISR operational criteria for a successful retrieval (Martonchik et al., 1998). Not surprisingly, the retrieved AODs do not compare favorably with the “true” values. Table 1 compares the retrieved spectral AOD for mixture 26 to the actual AODs for K2010 Case 12. Although the blue band AOD is estimated reasonably well, the spectral variation of AOD is much too steep.

These results indicated that the aerosol model used in K2010 does not have an adequate analog in the set of 74 mixtures used in the operational MISR aerosol retrieval. In an attempt to broaden the solution space of the LUT, additional tests were run using a richer set of aerosol mixtures containing

more types of non-absorbing aerosol particles. The pure particle set was supplemented with distributions having effective radii of  $0.57$  and  $1.28 \mu\text{m}$ . The new best fitting result, called “nonabsorbing 1,” is a mixture (by green band optical depth fraction) of 85% of a particle size distribution with an effective radius of  $0.57 \mu\text{m}$ , a  $\sigma$  of  $1.8$ , a refractive index of  $1.45-0.0i$ , and 15% of the same large particles used in mixture 26. The size parameters for this mixture as a whole are  $r_{\text{eff}}=0.98 \mu\text{m}$ ,  $v_{\text{eff}}=1.66$ . This combination provided an apparently good fit based on the (rather coarse) operational thresholds applied to the standard metrics. However, Fig. 3 shows that the percent differences in the BRFs predicted by the nonabsorbing 1 model and the synthetic data, while smaller than for mixture 26, are still large compared to the expected level of measurement uncertainty. The retrieved AOD errors for this model are spectrally flatter than for mixture 26, but are systematically too low by 12–16%, as shown in Table 1 for Case 12. Similar behavior is observed for the other K2010 cases. The overall curvature in the residuals in each spectral band as a function of scattering angle and the disagreement among all bands at scattering angles greater than  $150^\circ$  (see Fig. 3) suggest that this mixture still represents a non-optimal fit to the synthetic data.

### 3. Forward RT modeling

Before revisiting the blind aerosol retrieval exercise using an optimization approach instead of a LUT, we evaluated our forward RT modeling resources using as accuracy benchmarks the SCIATRAN-based radiances predicted for MISR viewing geometries in the K2010 study. The results presented below employ a successive-orders-of-scattering (SOS) 1D vector RT model for a coupled atmosphere–ocean system (Zhai et al., 2009), modified to allow the surface to be completely black. Two versions of this code are in use. SOS Version 1.0 is the original, highly flexible version of the code, and is used in batch mode for customized simulations, e.g., for plotting multiangle radiance curves that illustrate BRF sensitivity to aerosol or surface properties. SOS Version 2.0 is

**Table 1**

Spectral AOD for the mixture 26 and nonabsorbing 1 models obtained from the MISR LUT-based retrievals in K2010, compared to the “true” values for Case 12 ( $\tau_{412}=1.000$ ). Values in parentheses are the absolute difference between the retrieved and “true” AODs. Percent errors are shown at the bottom of each entry.

Data	443 nm	560 nm	670 nm	865 nm
K2010	1.001	0.994	0.978	0.935
Mixture 26	0.987	0.892	0.804	0.685
	(−0.014)	(−0.102)	(−0.174)	(−0.250)
	−1.40%	−10.26%	−17.79%	−26.74%
Nonabsorbing 1	0.846	0.863	0.857	0.807
	(−0.155)	(−0.131)	(−0.121)	(−0.128)
	−15.48%	−13.18%	−12.37%	−13.69%

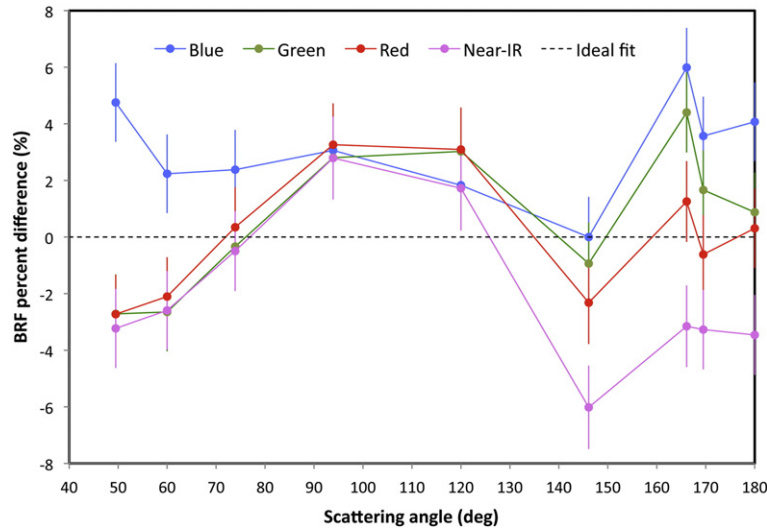


Fig. 3. Same as Fig. 2 but for the “nonabsorbing 1” mixture. Note the change in scale relative to Fig. 2.

a refactored (and slightly accelerated) version developed at JPL for ease of integration with other applications. The refactored code currently has less flexibility with respect to the vertical structure of the atmosphere, particularly with regard to how it handles the mixing of aerosol and molecular layers. On the other hand, it is well suited to repeated calls from an inversion scheme, as described in the following section.

Additionally, a 1D vector Monte Carlo model was developed by one of us (A.B. Davis) and F. Xu of JPL for high-accuracy benchmarking of RT in vertically-varying atmospheres. Here again, two variations of the Monte Carlo model are in use. The original version uses a *spatially continuous* representation of a Rayleigh-aerosol atmosphere where the molecular scattering coefficient varies exponentially with a given scale height and a uniform particulate layer of arbitrary thickness at any altitude. However, to rigorously compare the polarized radiances predicted by the Monte Carlo scheme with counterparts from deterministic 1D vector RT codes such as SCIATRAN or SOS, it proved necessary to develop a Monte Carlo variant that ingests the identical *spatially discrete* representation of the vertical structure and then enforces a homogeneity assumption inside each specified layer, irrespective of its optical thickness.

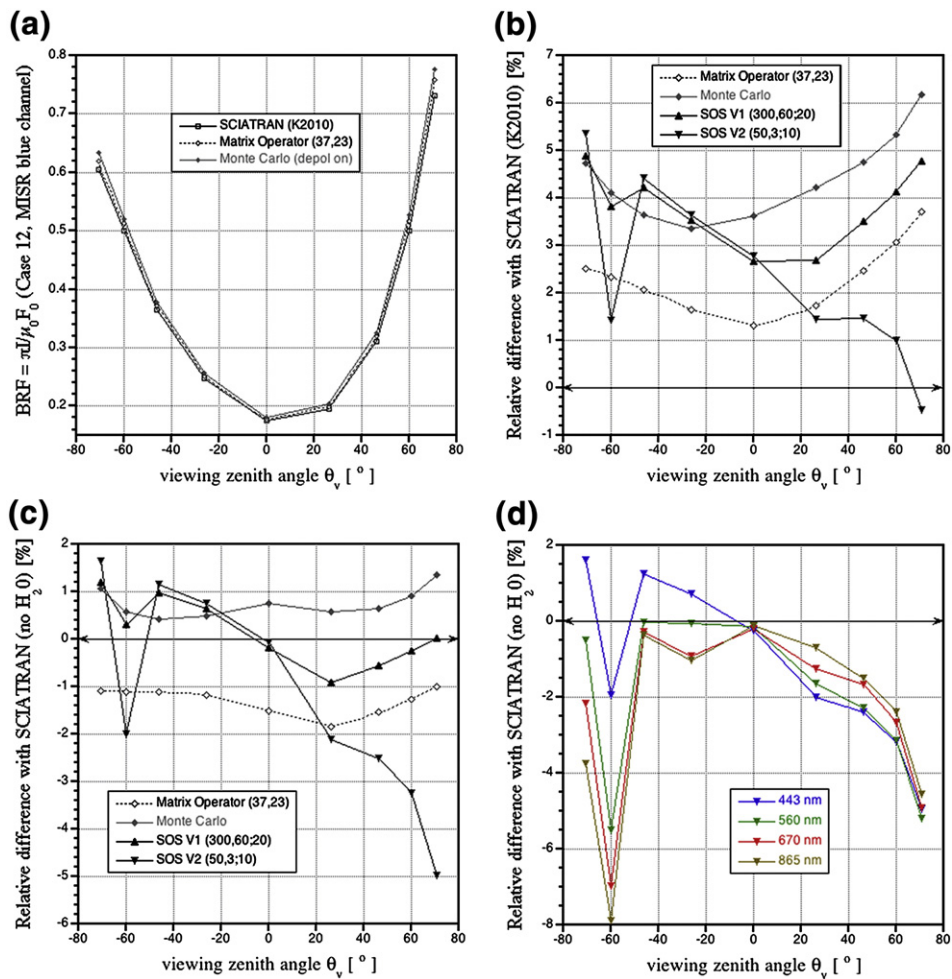
Fig. 4a shows the SCIATRAN BRFs for the 9 MISR angles used by K2010 for Case 12 ( $\tau_{412} = 1.0$ ) in the emulated blue channel (443 nm), where  $\tau_{443} = 1.00074$ . Also plotted are BRFs from: (1) the spatially discrete vector Monte Carlo model using the same parameters, including the same phase function tabulation used by SCIATRAN, the same Rayleigh depolarization, and the same vertical layering (in this case, three layers suffice, as defined in Table 2), and (2) the scalar Matrix Operator code, corrected for Rayleigh polarization effects. Fig. 4b plots residuals with respect to SCIATRAN radiances, and shows additional differences between SOS (with two different choices in the numbers of angular quadrature points  $N$ , azimuthal modes  $M$ , and scattering events  $S$ ) and SCIATRAN. In the SOS model, only two non-absorbing layers were used: either pure molecular scattering over pure aerosol scattering in V2.0; or pure

Rayleigh over a single mixed layer between 0 and 2 km in V1.0. For simplicity, Rayleigh depolarization effects were not incorporated into the input files of the SOS codes. In all cases, aerosol properties were computed from the particle definitions using the Mishchenko et al. (2002) Mie code.

The clear discrepancy between SCIATRAN and the other codes shown in Fig. 4b was traced to a single water vapor line that occurred at the 443 nm wavelength used to emulate MISR's blue channel. SCIATRAN was run in a line-by-line mode and a certain amount of water vapor was present but not identified a priori as a problem for any of the simulated spectral channels for any of the instruments involved in the K2010 study. Its presence at exactly 443 nm – the single wavelength used by K2010 to represent MISR's blue channel – was purely accidental. Fig. 4c is identical to Fig. 4b except that it uses updated SCIATRAN radiances without the inadvertent water vapor absorption. As shown in Fig. 4c, the Monte Carlo results now agree with corrected SCIATRAN predictions to better than 1%, which is a typical deviation when different treatments of the multiple scattering are compared, all other things being equal; see, e.g., Kokhanovsky et al. (2010b) for a recent benchmarking study. The same is true of SOS V1.0 when run with values of  $N$ ,  $M$ , and  $S$  chosen to ensure high accuracy, and with proper mixing of the Rayleigh and Mie scattering matrices in the 0–2 km layer. SOS V2.0 radiances still differ from SCIATRAN by several percent. By likely order of importance, the sources of this difference are: (a) coarser representation of the vertical profile; (b) lower-order angular discretizations (e.g.,  $M=3$ , compared with  $M=60$  in SOS V1.0); (c) different Mie single scattering computation details, such as truncation of the series expansion and the precise quadrature rule used to integrate over the particle size distribution; (d) SOS versus discrete-ordinate implementations of 1D vector RT (in particular, the use of fewer orders of scattering,  $S=10$ , compared with  $S=20$  in V1.0); and (e) lack of Rayleigh depolarization.

Since the use of a forward Monte Carlo model in a retrieval scheme is not a practical option, and upgrading the single-scattering and vertical profiling capability in SOS





**Fig. 4.** (a) Multi-angle SCIATRAN BRFs for Case 12 from the K2010 study at 443 nm compared to Monte Carlo and Matrix Operator results for the same atmosphere, except for a small amount of inadvertent water vapor absorption in SCIATRAN. (b) Monte Carlo- and Matrix Operator-SCIATRAN percent differentials from panel (a) supplemented by SOS-SCIATRAN differentials. SOS V1.0 used a 2-layer structure, the bottom (0–2 km) layer being an aerosol-Rayleigh mixture; V2.0 used a pure Rayleigh over pure aerosol profile. The numbers in parentheses for the deterministic codes ( $N, M; S$ ) are, respectively, the number of angular quadrature points, azimuthal modes, and in the case of SOS, number of scattering events. (c) Same as (b) but using updated SCIATRAN radiances without water vapor absorption. (d) Differentials between SOS V2.0 (at the lower accuracy settings used to speed up the computations) with respect to SCIATRAN (without water vapor absorption) for all four MISR bands.

V2.0 would not lead to any further insights into the aerosol remote sensing problem at hand, it was decided that the remainder of this study would be conducted by replacing SCIATRAN radiances with their SOS V2.0 counterparts for the same particle model and set of optical depth cases. The lower

values of  $N, M$ , and  $S$  (identified in Fig. 4b and c) were adopted for computational efficiency, and are acceptable for this study since the same configuration was used to create the synthetic data and to perform the inversions. Fig. 4d shows SOS V2.0 with these settings compared with SCIATRAN radiances (updated to eliminate the water vapor absorption) for all four MISR channels. We note that accuracy increases with wavelength in the forward-scattering plane and deteriorates at the  $-60^\circ$  view angle, which corresponds to the exact backscatter direction. This seems to be primarily due to truncation of the azimuthal Fourier series at  $M=3$ .

Ultimately, forward RT model accuracy and fidelity requirements should be formulated with measurement accuracy and precision in mind. We defer this kind of detailed end-to-end modeling process to future studies where such considerations are warranted, which is not the case in the present evaluation of the performance of the optimization approach to aerosol retrieval in an idealized simulation.

**Table 2**

Three-layer atmosphere used to benchmark forward 1D vector RT models, corresponding to Case 12 in K2010. From left to right: altitude  $z$  in km of the layer base, total (aerosol + Rayleigh) optical depth  $\tau_{443}$  (from TOA) of layer base, optical thickness  $\Delta\tau$  of the layer, Rayleigh fraction in scattering phase function, and single scattering albedo (SSA) of the layer.

$z_{\text{base}}$ (km)	$\tau$ at 443 nm (to TOA)	Layer optical thickness $\Delta\tau$	Rayleigh fraction	SSA
2.0	0.185043	0.185043	1	1.0
1.0	0.710051	0.525008	0.0469292	1.0
0.0	1.238340	0.528289	0.0528474	1.0

#### 4. Optimization approach

In this section, we frame the aerosol retrieval problem from modeled radiances in terms of continuous optimization. In general, the objective of optimization is to find a set of parameters  $\mathbf{x}$  that minimizes a cost function  $C(\mathbf{x})$ . Algorithms to solve this problem are typically iterative: given an initial guess  $\mathbf{x}_0$ , the algorithm produces a sequence  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K\}$  of improved estimates, or *iterates*, which (ideally) converge to the correct solution. The notation  $\mathbf{x}_k$  indicates the  $k$ th iteration on  $\mathbf{x}$ , and  $K$  indicates the last iteration. The strategy for determining the next iterate in the sequence is a distinguishing characteristic of different algorithms (Nocedal and Wright, 2006).

To relate this methodology to our problem, we first identify the aerosol model parameters to retrieve. The five parameters selected are listed in Table 3, where the first four characterize the aerosol phase function and are inputs to the Mie particle scattering calculation, and the last is the AOD at a reference wavelength of 550 nm. For each parameter, we restrict the solution to fall within some interval, defining a rectangular box in the five-dimensional solution space. In optimization terminology, these intervals are aptly called *box constraints*. They are chosen here to represent a plausible range of aerosol properties observed over the Earth. Although each interval is independently specified, some regions of this five-dimensional box correspond to aerosol models not normally found in nature, e.g., those corresponding to particles with single-scattering albedo (SSA) of less than 0.5. In future work, an improved setup might explicitly constrain the SSA using Lagrange multipliers. Note that we use  $10^{-4}$  as a proxy for zero AOD because extremely small values relative to the Rayleigh optical depth slow down the RT calculation considerably, due to the way the integration interval size is chosen in SOS when integrating with respect to optical depth (Zhai et al., 2009).

Given some point  $\mathbf{x}$  in the parameter space, the phase function is calculated for each of the four MISR spectral bands, followed by the four AODs, determined using the ratio of each band's extinction cross section to that of the reference:

$$\tau_\lambda = \tau_{550} \frac{\kappa_\lambda}{\kappa_{550}}. \quad (2)$$

The 36 BRFs are then calculated and compared with the synthetic MISR observations, generated beforehand using the

same forward model (SOS V2.0). There are many reasonable choices for the cost function, each with different benefits and implications. For simplicity, we chose an unweighted least-squares cost function:

$$C(\mathbf{x}) = \|\mathbf{C}(\mathbf{x})\|^2 = \sum_{j=1}^{36} [R_j^{model}(\mathbf{x}) - R_j^{obs}]^2 \quad (3)$$

where  $R_j^{model}(\mathbf{x})$  is the modeled BRF in channel (4 bands  $\times$  9 angles)  $j$ , and  $R_j^{obs}$  is the simulated observation in channel  $j$ . It is possible that other cost functions may be more robust or efficient. One example of an alternative cost function uses the  $L^1$  norm, in which the squared difference in Eq. (3) is replaced by the absolute value of the difference. This alternative has steeper gradients near the solution, which can provide faster convergence, but is non-differentiable at the solution, which might violate the assumptions of a chosen optimization algorithm. One might additionally scale the least squares cost function by some measurement covariance matrix (Rodgers, 2000), or include a second term that penalizes the departure from some a priori solution. The latter approach is formally known as *optimal estimation*, and is useful when the system is underdetermined. This technique has been applied to aerosol remote sensing, e.g., in a coupled retrieval of atmospheric and surface reflectance parameters from the Meteosat Second Generation satellite (Govaerts et al., 2010) and in retrieval of aerosol properties using the airborne Research Scanning Polarimeter (Waquet et al., 2009). Dubovik et al. (2010) developed an optimized inversion approach that solves for a large number of parameters describing the AOD, size distribution, particle optical properties, and reflection from the underlying surface. Their approach invokes constraints on smoothness of the size distribution and on the spectral variation of the real and imaginary refractive indices and surface parameters. Good results were obtained in case studies with POLDER data.

In the optimization experiments described in this paper, we adopted a monomodal lognormal aerosol size distribution and constrained the surface to be black, but purposely did not use any a priori parameter solution nor explicitly add noise to the synthetic data in order to emulate the blind retrieval setup of the original K2010 study. The Levenberg–Marquardt algorithm was employed for parameter optimization. While alternative algorithms can be used, this method suited our objective of using a simple optimization framework to test the effectiveness of searching a continuous solution space. Benchmarking tests also showed Levenberg–Marquardt optimization to be more efficient than two other approaches that were considered (a downhill simplex method and a genetic algorithm), while retaining a comparable level of accuracy.

At each iteration, the Levenberg–Marquardt algorithm computes the direction to move in parameter space for the next iteration by minimizing a quadratic approximation of the cost function in a region (called the *trust region*) around the current iterate, where the radius of this region is governed by the *damping parameter*,  $\mu_k$ , for iteration  $k$ . The specific implementation we employed is outlined in Algorithm 3.12 of Kanzow et al. (2004). Levenberg–Marquardt is often described as a combination of the Gauss–Newton and Gradient Descent methods; when far from the solution, it acts like Gradient Descent, whereas near the solution, the iteration proceeds

**Table 3**

Parameters fitted using optimization. The third and fourth columns show the allowed range (inclusive) of retrieved solutions, corresponding to the box constraints imposed on the overall solution space. The “true” values of the parameters correspond to the K2010 aerosol model used to calculate the synthetic TOA radiances, as described in Section 2.

Parameter	Symbol	Min	Max	True
Median particle radius ( $\mu\text{m}$ )	$r_m$	0.02	1.00	0.1
Antilogarithm of size distribution standard deviation	$\sigma$	1.1	3.0	2.718 ( $e$ )
Real part of the refractive index	$n_r$	1.3	1.7	1.38
Imaginary part of the refractive index	$n_i$	0.0	0.2	0.0
Reference AOD	$\tau_{550}$	0.0001	7.0	0.0–5.0

closer to the Gauss–Newton direction (Lourakis, 2004), governed by the damping parameter. This is equivalent to solving the following linear system:

$$(\mathbf{J}_k^T \mathbf{J}_k + \mu_k \mathbf{I}) \mathbf{d}_k = -\mathbf{J}_k^T C(\mathbf{x}_k), \quad \mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k. \quad (4)$$

Here the matrix  $\mathbf{J}_k$  denotes the Jacobian of  $R_j^{model}(\mathbf{x}_k)$  and  $\mathbf{I}$  is the identity matrix. During a given iteration, several values of the damping parameter might be tried, until one is found which sufficiently reduces the cost function. In this way, the damping parameter helps navigate highly nonlinear regions of the parameter space, adaptively adjusting the trust region radius until the quadratic approximation fits the cost function well enough to produce a good descent direction. The algorithm handles the box constraints simply by computing the direction without constraints, and then if necessary, projecting the new estimate  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}_k$  to the nearest point on the box. In practice, we find that solutions sometimes get “stuck” on the box boundary, leading to non-convergence. Further investigation is necessary to determine the reason for this behavior.

The implementation of the Levenberg–Marquardt algorithm that we employed terminates if any of the following quantities drop below a user-defined threshold: (a) the cost function (also referred to in the following discussion as the “residual”), (b) the gradient of the cost function vector, or (c) the step length ( $\mathbf{d}_k$ ) from one iteration to the next. These thresholds and other parameters used to operate the algorithm are listed in Table 4. The cost function threshold corresponds roughly to the value implied by MISR’s channel-to-channel noise levels. In order to simplify diagnostics, the gradient threshold was assigned a tiny value so that the retrievals would not terminate due to a small gradient. Both the maximum number of iterations and the minimum relative step-length were determined by trial and error and the parameter perturbation scheme for the Jacobian computation was kept at its default as defined in the software package. Various other parameters used the values listed in Kanzow et al. (2004). In the following results, we consider a retrieval to be successful only if the algorithm terminated due to a small residual.

## 5. Retrieval results

Synthetic, noiseless data were generated using SOS V2.0 and the same aerosol model and set of 16 AODs as in the K2010 study. Parameters affecting accuracy of the forward model calculation (e.g., number of quadrature points) were

set to identical levels for both the generation of synthetic data and in the inversions, as discussed in Section 3.

One important consideration when using many optimization algorithms, including Levenberg–Marquardt, is that the retrieved solution is not guaranteed to be the global minimum of the cost function; it may instead only be a local solution. In fact, there is often no feasible way to determine whether a particular solution is truly the global minimum. Thus, if the cost function is degenerate or has multiple local minima, the solution will be sensitive to initial conditions. An operational retrieval algorithm should thus initialize as close to the optimal solution as possible, which might be accomplished using, e.g., climatological data, a retrieved solution from a nearby pixel, the output of a LUT search, or some combination of these. For the present study, we dealt with this issue by starting the optimization at 10 different points in parameter space for each of the 16 cases. The same 10 starting points were used in all 16 cases. The starting points were chosen randomly in all parameters with the exception of  $n_i$ , which was initialized to 0, and  $\tau_{550}$ , which was initialized to 0.001. The initialization of  $n_i$  is reasonable since most aerosols observed over the Earth are either non-absorbing or very nearly so. Starting  $\tau_{550}$  at 0.001 worked better than random initialization, as we found that when  $\tau_{550}$  was initialized to a value greater than the “truth,” the retrievals tended to wander further from the solution, increasing  $\tau_{550}$  and  $n_i$  and falling into a well-known degeneracy between AOD and SSA. In contrast, when  $\tau_{550}$  was initialized to a value below the “truth,” this degeneracy did not adversely impact the optimization. Furthermore, starting off at a low value of  $\tau_{550}$  did not cause problems even when retrieving the highest AODs in this experiment. Table 5 shows the 10 randomly chosen initial values for the other three parameters.

We explored two different algorithms for the retrievals—a straightforward 1-step optimization, and a 2-step approach. In the 1-step procedure, full multiple scattering RT calculations are performed at each iteration. In the 2-step procedure, the first step consists of an optimization initialized as described above, but computes only the first order of scattering (numerically, still using SOS V2.0), and terminates based upon less strict criteria than in the final step (see Table 4). The result of this step then becomes the starting point for the second optimization step, which employs the full multiple-scattering calculations.

In the following discussion, we use the terms “convergence” or “success” to indicate trials for which the cost function has dropped below the prespecified threshold of

**Table 4**

Parameters controlling the operation of the Levenberg–Marquardt algorithm. Differences in stopping thresholds between the 1-step procedure and the second part of the 2-step arise from the fact that the initial aerosol parameters in the 1-step are typically farther from the solution. Because the first part of the 2-step procedure is essentially a preprocessing step, thresholds are less strict than in the second part.

Algorithm parameter	1-step	2-step, part 1	2-step, part 2
Stopping threshold: cost function: $C(\mathbf{x}_k)$	$10^{-6}$	$10^{-4}$	$10^{-6}$
Stopping threshold: magnitude of gradient of cost function vector: $\ \nabla C(\mathbf{x}_k)\  = \ \mathbf{J}_k^T C(\mathbf{x}_k)\ $	$10^{-15}$	$10^{-15}$	$10^{-15}$
Stopping threshold: step length to next iterate relative to current iterate: $\ \mathbf{d}_k\ /\ \mathbf{x}_k\ $	$10^{-4}$	$10^{-3}$	$10^{-3}$
Maximum number of iterations	300	20	50
Parameter perturbation increment for finite-difference Jacobian calculation	$\max(10^{-4}\mathbf{x}_k, 10^{-6})$	$\max(10^{-4}\mathbf{x}_k, 10^{-6})$	$\max(10^{-4}\mathbf{x}_k, 10^{-6})$



**Table 5**

Initial guesses were chosen randomly for median radius, antilogarithm of the size distribution standard deviation, and real refractive index, using the same 10 points for all 16 test cases.

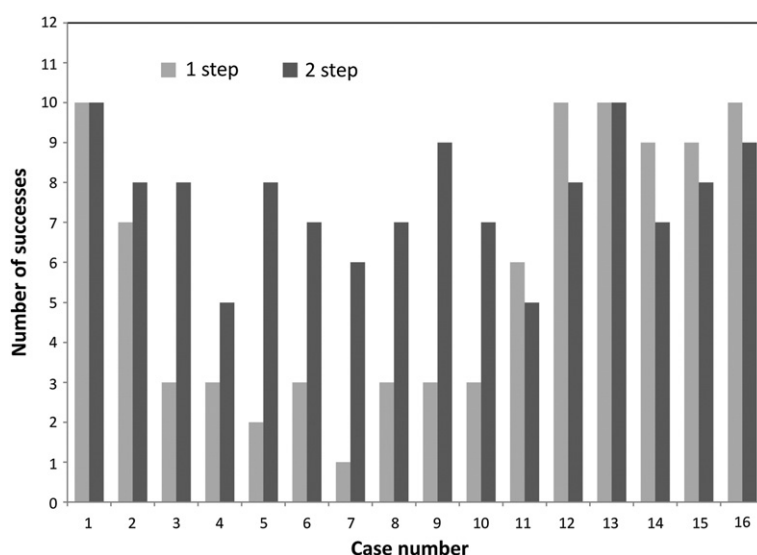
Parameter	Trial number									
	1	2	3	4	5	6	7	8	9	10
$r_m$	0.563	0.615	0.234	0.776	0.950	0.082	0.674	0.367	0.693	0.913
$\sigma$	2.007	1.131	2.075	2.101	2.099	1.553	2.110	2.407	2.917	2.017
$n_r$	1.552	1.397	1.636	1.608	1.334	1.688	1.316	1.366	1.535	1.386

$10^{-6}$ . These terms do not necessarily imply that the exact global minimum has been reached. Rather, a value below the threshold indicates that the resulting model BRFs are not distinguishable from the synthetic data to within the assumed level of measurement uncertainty (Bruegge et al., 1998, 2002). Both the 1-step and 2-step retrieval approaches converged at least once out of the 10 initialization trials for each of the 16 cases. Fig. 5 shows the number of successes as a function of case number for both the 1-step and 2-step procedures. In terms of the number of successful convergences obtained, the 2-step algorithm performed on par or better than the 1-step in Cases 1–10, with the opposite result in Cases 11–16. The transition between these two regimes, at Case 11, occurs at  $\tau_{412} = 0.5$ . When not specified, the remaining figures and discussion all correspond to the 2-step approach.

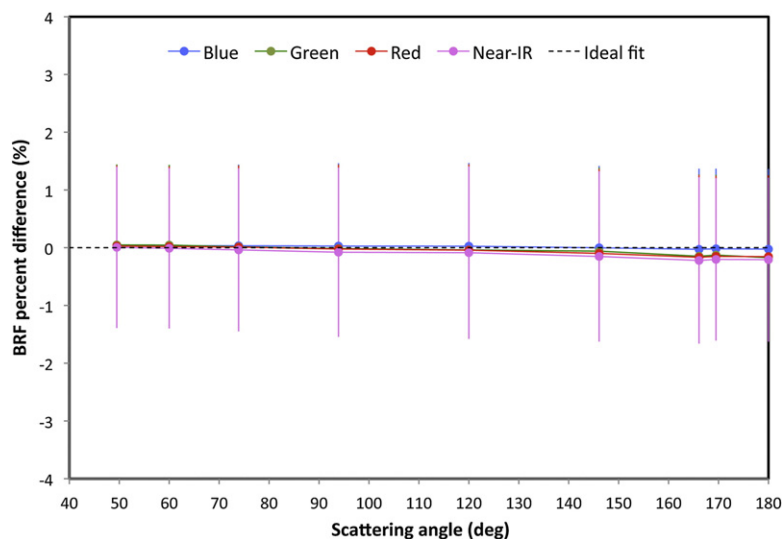
Fig. 6 shows BRF percent differences from each of the 36 synthetic measurements for the retrieval in Case 12 having the largest value of cost function that met the convergence criterion (i.e., the “worst-case” successful retrieval). Comparison of this figure with Figs. 2 and 3 demonstrates the significant advantage of the optimization approach relative to the LUT calculations that were used in the K2010 study. The retrieval results for Cases 2–16 are summarized in Figs. 7 and 8, which show the mean and standard deviation of each of the retrieved parameters (AOD relative to its true value in Fig. 7 and  $r_m$ ,  $\sigma$ ,  $n_r$ ,  $n_i$  in Fig. 8), computed from the set of trials which

successfully converged using the 2-step approach. (Particle properties are not relevant for Case 1 since AOD = 0, and all 10 trials in Case 1 correctly converged to negligible AOD.) There is a clear trend demonstrated by all five parameters showing both increasing accuracy and reduced variance among the different trials with increasing AOD. The AOD retrievals achieve almost zero bias starting with Case 5 ( $\tau_{412} = 0.04$ ), with all parameters converging in a similar fashion by Case 8 ( $\tau_{412} = 0.2$ ). This result is consistent with the conclusion of Kahn et al. (2010), who note reduced sensitivity of MISR to particle properties when mid-visible AOD is  $< 0.15$ . Interestingly, Fig. 7 shows that even at low aerosol loading, the bias and uncertainty in AOD are low ( $< 10\%$  of the AOD), despite the departure of the particle properties from their true values. Much (though not all) of the spread in retrieved AOD at low aerosol loading is due to departure of the SSA from unity, consistent with the fact that in the single scattering regime AOD and SSA are not uniquely separable. Using the Matrix Operator code to calculate the cost function in a hyperspace of simulated measurements  $R^{\text{mod } el}(\tau, r_m, \sigma, n_r, n_i)$  confirms this expectation (see Fig. 9a). Fig. 9b shows that at high AOD, where multiple scattering plays a major role, the cost function increases rapidly when the aerosol departs from the true values of AOD and  $n_i$ .

Other degeneracies are also apparent at low AOD, notably a rough inverse correlation between  $r_m$  and  $\sigma$ . In Cases 2–4 ( $\tau_{412} = 0.01 - 0.03$ ),  $r_m$  varied from a high of  $0.14 \mu\text{m}$  to a low



**Fig. 5.** Number of retrieval successes (i.e., convergence with residuals below the specified threshold) out of 10 initialization trials as a function of K2010 case number for the 1-step and 2-step algorithms.

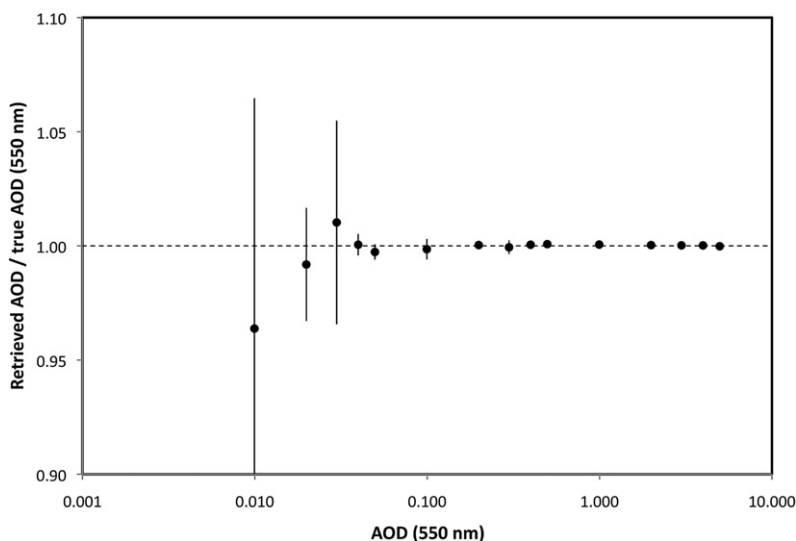


**Fig. 6.** Same as Figs. 2 and 3 but for a successful aerosol model determined using optimization approach. Results are shown for the worst fitting model that met the cost function convergence criterion. Note the change in scale relative to Figs. 2 and 3.

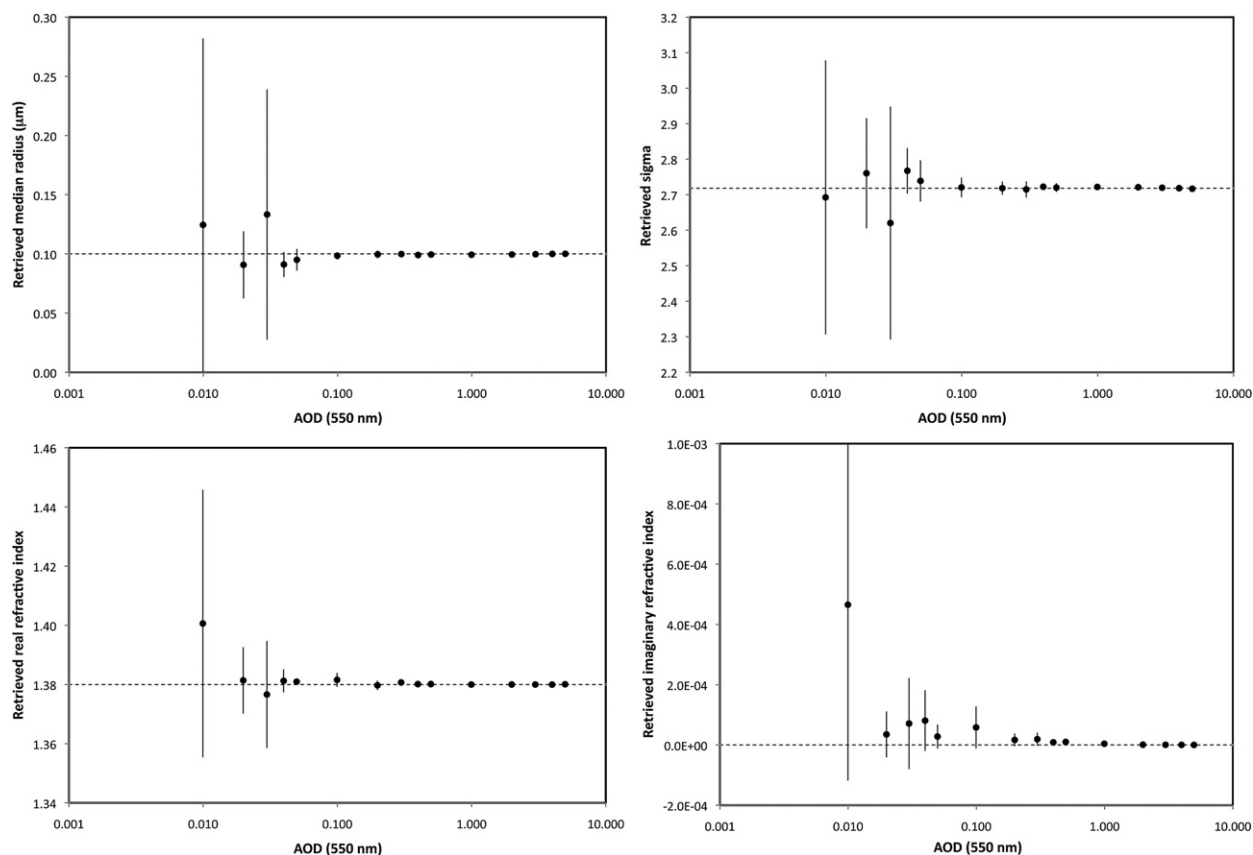
of  $0.05 \mu\text{m}$  for the successful trials, whereas  $\sigma$  varied from 2.5 to 3.0. A contour plot of the cost function generated using the Matrix Operator code confirms that in Case 4, for example, the residuals over this range of  $r_m$  and  $\sigma$  fall below or approach our convergence criterion, when all other parameters are fixed at their true values (Fig. 9c). The diminished error bars in Fig. 8 indicate that this degeneracy is overcome when AOD exceeds 0.1, and Fig. 9d shows that at high AOD the cost function is large except in the immediate vicinity of the correct solution.

An interesting trial that illuminates the inverse correlation of  $r_m$  and  $\sigma$  at low AOD is one from Case 4, in which the cost function  $C = 6 \times 10^{-6}$ , and the largest residuals between the model BRF and synthetic observations are just above the level of measurement uncertainty. In this “nearly converged” example,

the retrieved  $r_m$  was  $0.35 \mu\text{m}$ ,  $\sigma$  was 1.98, and the retrieved AOD was 0.033 (i.e., within 10% of the correct value of 0.03). Fig. 10 compares the lognormal size distribution for this retrieval with the results from a converged trial where  $r_m$  was  $0.08 \mu\text{m}$  and  $\sigma$  was 2.779 (i.e., much closer to the true values of  $r_m = 0.1 \mu\text{m}$  and  $\sigma = 2.718$ ). In the “nearly converged” case,  $r_{\text{eff}} = 1.12 \mu\text{m}$ , compared to  $1.09 \mu\text{m}$  for the converged case. This similarity in  $r_{\text{eff}}$ , despite the difference in shape of the size distribution, leads to very similar phase functions and spectral extinction, hence the low residual in both cases. In contrast, the bottom panels in Fig. 10 show a trial from Case 4 in which  $r_m$  was  $0.05 \mu\text{m}$  and  $\sigma$  was 2.921 ( $r_{\text{eff}} = 0.91 \mu\text{m}$ ), but with a final cost function of  $4 \times 10^{-4}$  (the next lowest value after  $6 \times 10^{-6}$  among the 10 trials), clearly indicating non-convergence. The large residual is most likely due to the fact that the retrieved value of  $n_i$  was 0.05,



**Fig. 7.** Mean value of retrieved AOD (at 550 nm) for the successfully converged results in K2010 Cases 2–16 divided by the “true” value. Error bars indicate one standard deviation. The dashed line shows the ideal value of 1.0.



**Fig. 8.** Retrieved particle properties as a function of optical depth for the successfully converged results in K2010 Cases 2–16. The means among all successful trials are shown as the dots, and the error bars indicate one standard deviation. Clockwise from upper left, the four panels show results for  $r_m$  (the true value of  $0.10 \mu\text{m}$  is indicated by the dashed line),  $\sigma$  (true value = 2.718), imaginary part of refractive index  $n_i$  (true value = 0), and real part of refractive index  $n_r$  (true value = 1.38).

corresponding to an unreasonably low SSA of 0.65. To compensate for the increased absorption, the retrieved AOD was 0.061. For this trial, the spectral phase functions do not resemble the truth very well, and show suppressed backscatter (see Fig. 10). Even at this low value of AOD, the relatively high value for the cost function and departure of the modeled BRFs from the input values (differing in the near-infrared by as much as 17% at large scattering angles) caused this solution to be rejected.

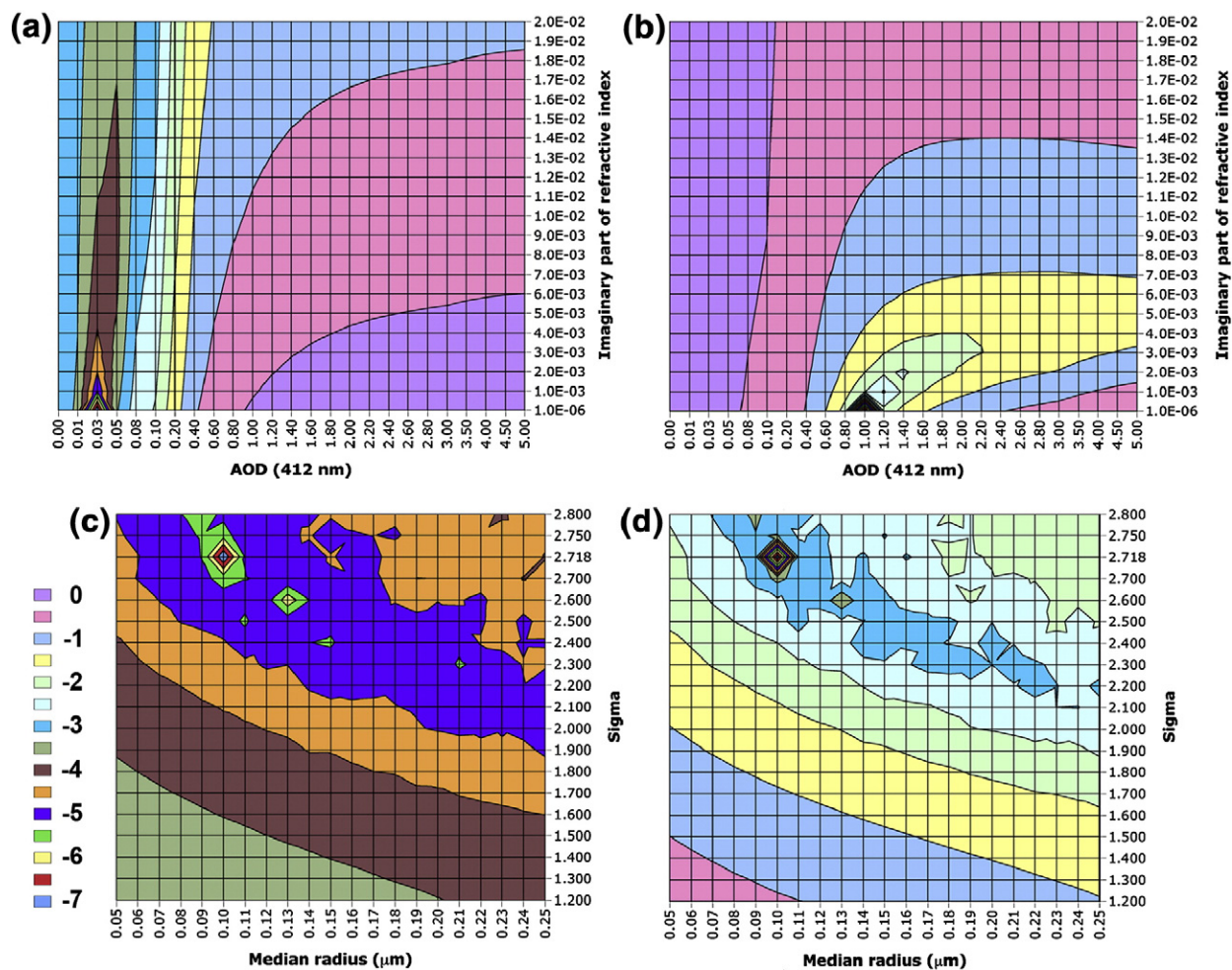
## 6. Conclusions

The results of this investigation show that, at least for the idealized scenario posed in the original K2010 study, retrievals based on least-squares optimization of continuously varying aerosol parameters using multiangle radiances yield far superior results compared to those obtained with a LUT. Undoubtedly, using a LUT containing the “correct” aerosol model would also yield excellent results, but the advantage of the optimization approach is that it obviates the need to discretely predefine the solution space. The limited extent of the MISR LUT was the principal cause of the AOD biases found in the K2010 paper; that is, the size distribution of the target aerosol was unlike any distribution contained in the LUT.

The K2010 study found that the LUT-based MISR retrieval ranked higher than many other approaches, but was surpassed

in accuracy by an optimization-based retrieval using simulated multiangle, spectropolarimetric observations. We conclude here that limitations of the MISR LUT, and not the lack of polarimetric information in the simulated MISR data, account for this result. Observational and theoretical studies have shown that polarimetry adds significant value to aerosol retrievals (e.g., Chowdhary et al., 2002; Hasekamp and Landgraf, 2007; Lebsock et al., 2007) and those conclusions are not disputed here. Indeed, the degeneracy between  $r_m$  and  $\sigma$  at low AOD for intensity-only observations can potentially be mitigated through the use of polarimetry (Chowdhary et al., 2001). Our findings suggest, however, that a reliance on discretized LUTs to interpret multiangle spectral intensity-only observations may not take full advantage of the information content of such measurements.

The optimization calculations described in this paper were performed using sequential runs on an Intel Xeon X5450 3.0 GHz processor with 16 GB RAM. Table 6 shows the average runtime, number of iterations, and number of cost function evaluations, averaged over all successful retrievals per case, and then averaged over all cases. The individual metrics increase by roughly 50% from Case 2 to Case 16. Although Table 6 shows the 2-step process to be slightly slower, it significantly reduces the sensitivity to initial conditions, as shown in Fig. 5. The current runtime results are unacceptable for a global aerosol imager such as MISR, so



**Fig. 9.** Example contour plots of the logarithm of the cost function. In each panel  $\log_{10}(C)$  is plotted as a function of the indicated aerosol parameters (other aerosol model parameters are fixed at their “true” values). The scale shown in the legend at lower left applies to all four panels. (a) Cost function vs.  $\tau_{412}$  and  $n_i$  for simulation Case 4 ( $\tau_{412} = 0.03$ ), showing how increased AOD compensates for greater aerosol absorption. (b) Same as (a) except for Case 12 ( $\tau_{412} = 1.00$ ). (c) Cost function vs.  $r_m$  and  $\sigma$  for Case 4 showing the inverse correlation between these variables. (d) Same as (c) except for Case 12, demonstrating a similar inverse relationship except that the cost function takes on much higher values away from the location where the parameters take on their true values ( $r_m = 0.10 \mu\text{m}$ ,  $\sigma = 2.718$ ). At the exact solution  $C$  is identically zero, but for the purpose of generating these plots, we set  $\log_{10}(C)$  to  $-7$  at this location.

methods of speeding up the computations (e.g., code parallelization, the use of Graphics Processing Units, or algorithmic and coding shortcuts) are clearly needed in order to make such an approach viable in an operational environment. The current computations used finite differences in the calculation of the Jacobian matrix. We have been exploring RT code linearization as a means of reducing computation time and increasing accuracy (see, e.g., Hasekamp and Landgraf, 2005); implementation within our Matrix Operator code reduces runtimes by about a factor of three. Since the ultimate implementation of aerosol retrieval hinges greatly on the time required, this approach offers a potentially significant advantage to MISR, as well as to advanced follow-on instruments currently in development, such as the Multiangle SpectroPolarimetric Imager (MSPI) (Diner et al., 2010).

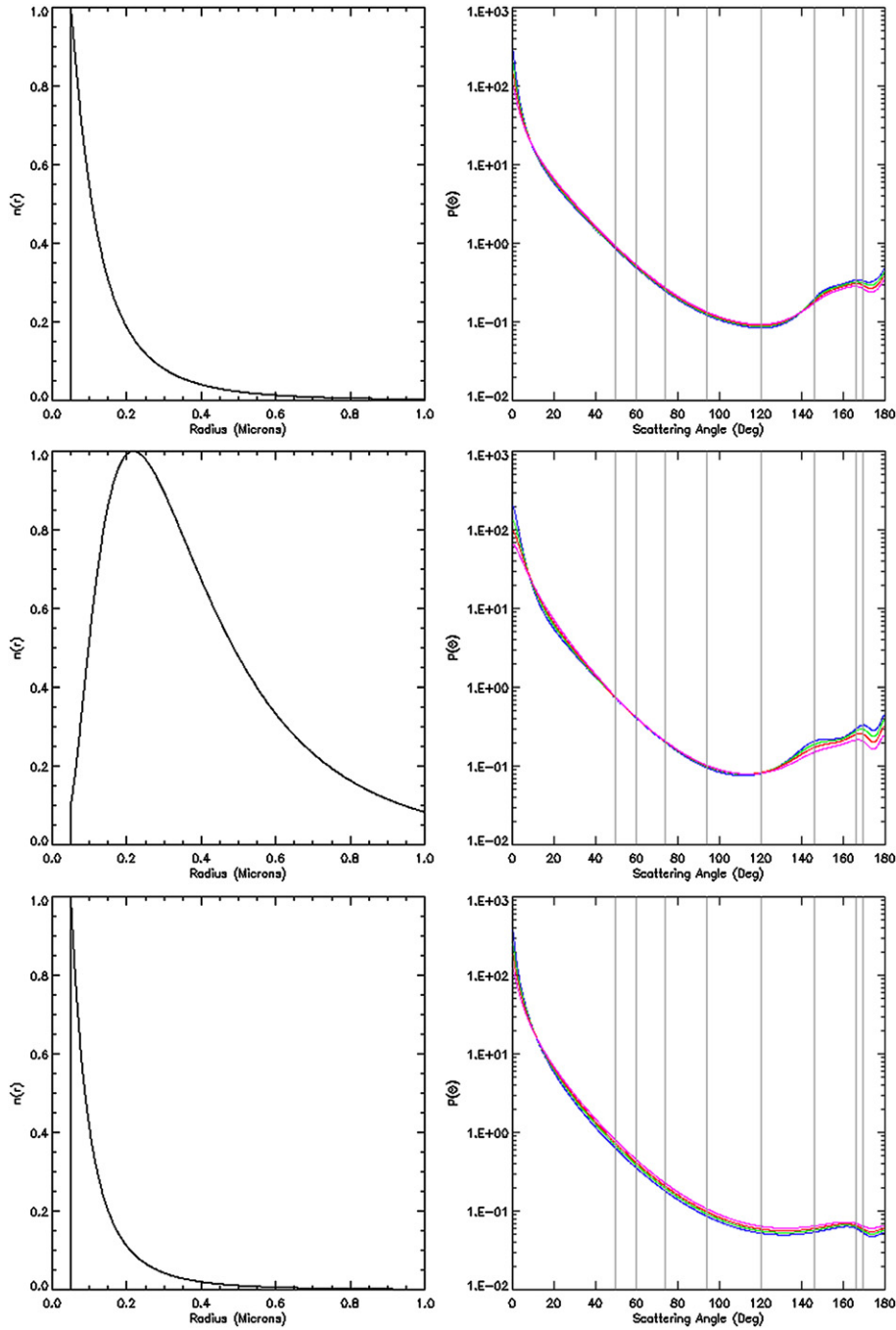
For consistency with the strategy employed in the original K2010 study, measurement noise was not included in the

retrievals discussed here, other than to use the magnitude of MISR measurement uncertainties as a point of reference for assessing goodness of fit. Nonetheless, the impact of measurement noise on retrieval sensitivity needs to be examined in detail. We also plan to extend the present study to handle broader spectral coverage, polarimetric measurements, and more general atmospheric and surface conditions. Along with development of computationally efficient methods of traversing the aerosol solution space and application of the optimization approach to actual MISR data, these topics are high-priority avenues for further investigation.

### Acknowledgments

We are grateful to several JPL colleagues for their assistance: Mike Smyth, for refactoring the SOS RT code;





**Fig. 10.** Retrieved size distributions (left) and spectral phase functions (right) for three trials in Case 4. The top panels correspond to the trial with the largest cost function  $C$  below the threshold of  $10^{-6}$ ; the middle panels to a "nearly converged" case with  $C = 6 \times 10^{-6}$ ; and the bottom panels to a trial with the smallest value of  $C$  ( $4 \times 10^{-4}$ ) that clearly did not converge.

**Table 6**

Retrieval statistics averaged over all successful retrievals per case and then over all cases.

	# cost function evaluations	Runtime (min)	# iterations
1-step	138	23	14
2-step	186	27	19

Seungwon Lee for help with the optimizations; and Robert West for valuable advice. We thank Yoav Schechner of the Technion Israel Institute of Technology for suggesting the 2-step approach. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA).

## References

- Abdou, W.A., Martonchik, J.V., Kahn, R.A., West, R.A., Diner, D.J., 1997. A modified linear-mixing method for calculating atmospheric path radiances of aerosol mixtures. *J. Geophys. Res.* 102, 16883–16888. doi:10.1029/96JD03434.
- Bruegge, C.J., Chrien, N.L., Diner, D.J., Kahn, R.A., Martonchik, J.V., 1998. MISR radiometric uncertainty analyses and their utilization within geophysical retrievals. *Metrologia* 35, 571–579.
- Bruegge, C.J., Chrien, N.L., Ando, R.R., Diner, D.J., Abdou, W.A., Helmlinger, M.C., Pilorz, S.H., Thome, K.J., 2002. Early validation of the Multi-angle Imaging SpectroRadiometer (MISR) radiometric scale. *IEEE Trans. Geosci. Remote Sens.* 40, 1477–1492.
- Chowdhary, J., Cairns, B., Mishchenko, M., Travis, L., 2001. Retrieval of aerosol properties over the ocean using multispectral and multiangle photopolarimetric measurements from the Research Scanning Polarimeter. *Geophys. Res. Lett.* 28, 243–246.
- Chowdhary, J., Cairns, B., Travis, L.D., 2002. Case studies of aerosol retrievals over the ocean from multiangle, multispectral photopolarimetric remote sensing data. *J. Atmos. Sci.* 59, 383–397.
- Diner, D.J., Beckert, J.C., Reilly, T.H., Bruegge, C.J., Conel, J.E., Kahn, R., Martonchik, J.V., Ackerman, T.P., Davies, R., Gerstl, S.A.W., Gordon, H.R., Muller, J.-P., Myneni, R., Sellers, P.J., Pinty, B., Verstraete, M.M., 1998. Multi-angle Imaging Spectroradiometer (MISR) instrument description and experiment overview. *IEEE Trans. Geosci. Remote Sens.* 36, 1072–1087.
- Diner, D.J., Davis, A., Hancock, B., Geier, S., Rheingans, B., Jovanovic, V., Bull, M., Rider, D.M., Chipman, R.A., Mahler, A., McClain, S.C., 2010. First results from a dual photoelastic modulator-based polarimetric camera. *Appl. Opt.* 49, 2929–2946.
- Dubovik, O., Herman, M., Holdak, A., Lapyonok, T., Tanré, D., Deuzé, J.L., Ducos, F., Sinyuk, A., Lopatin, A., 2010. Statistically optimized inversion algorithm for enhanced retrieval of aerosol properties from spectral multi-angle polarimetric satellite observations. *Atmos. Meas. Tech. Discuss.* 3, 4967–5077.
- Govaerts, Y.M., Wagner, S., Lattanzio, A., Watts, P., 2010. Joint retrieval of surface reflectance and aerosol optical depth from MSG/SEVIRI observations with an optimal estimation approach: 1 Theory. *J. Geophys. Res.* 115, D02203. doi:10.1029/2009JD011779.
- Grant, I.P., Hunt, G.E., 1968. Solution of radiative transfer problems using the invariant Sn method. *Mon. Not. R. Astron. Soc.* 141, 27–41.
- Hasekamp, O.P., Landgraf, J., 2005. Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing. *J. Geophys. Res.* 110, D04203. doi:10.1029/2004JD005260.
- Hasekamp, O.P., Landgraf, J., 2007. Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements. *Appl. Opt.* 46, 3332–3344.
- Kahn, R., Banerjee, P., McDonald, D., 2001. Sensitivity of multiangle imaging to natural mixtures of aerosols over ocean. *J. Geophys. Res.* 106, 18219–18238.
- Kahn, R.A., Garay, M.J., Nelson, D.L., Yau, K.K., Bull, M.A., Gaitley, B.J., Martonchik, J.V., Levy, R.C., 2007. Satellite-derived aerosol optical depth over dark water from MISR and MODIS: comparisons with AERONET and implications for climatological studies. *J. Geophys. Res.* 112, D18205. doi:10.1029/2006JD008175.
- Kahn, R.A., Gaitley, B.J., Garay, M.J., Diner, D.J., Eck, T.F., Smirnov, A., Holben, B.N., 2010. MISR global aerosol product assessment by comparison with AERONET. *J. Geophys. Res.* 115, D23209. doi:10.1029/2010JD014601.
- Kanzow, C., Yamashita, N., Fukushima, M., 2004. Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints. *J. Comput. Appl. Math.* 172, 375–397.
- Katsev, I.L., Prikhach, A.S., Zege, E.P., Ivanov, A.P., Kokhanovsky, A.A., 2009. Iterative procedure for retrieval of spectral aerosol optical thickness and surface reflectance from satellite data using fast radiative transfer code and its application to MERIS measurements. In: Kokhanovsky, A.A., de Leeuw, G. (Eds.), *Satellite Aerosol Remote Sensing Over Land*. Springer, Berlin.
- Kokhanovsky, A.A., Breon, F.-M., Cacciari, A., Carboni, E., Diner, D., Di Nicolantonio, W., Grainger, R.G., Grey, W.M.F., Höller, R., Lee, K.-H., Li, Z., North, P.R.J., Sayer, A.M., Thomas, G.E., von Hoyningen-Huene, W., 2007. Aerosol remote sensing over land: a comparison of satellite retrievals using different algorithms and instruments. *Atmos. Res.* 85, 372–394.
- Kokhanovsky, A.A., Deuzé, J.L., Diner, D.J., Dubovik, O., Ducos, F., Emde, C., Garay, M.J., Grainger, R.G., Heckel, A., Herman, M., Katsev, I.L., Keller, J., Levy, R., North, P.R.J., Prikhach, A.S., Rozanov, V.V., Sayer, A.M., Ota, Y., Tanré, D., Thomas, G.E., Zege, E.P., 2010a. The inter-comparison of major satellite aerosol retrieval algorithms using simulated intensity and polarization characteristics of reflected light. *Atmos. Meas. Tech.* 3, 909–932.
- Kokhanovsky, A.A., Budak, V.P., Cornet, C., Duan, M., Emde, C., Katsev, I.L., Klyukov, D.A., Korkin, S.V., Labonnote, L.C., Mayer, B., Min, Q., Nakajima, T., Ota, Y., Prikhach, A.S., Rozanov, V.V., Yokota, T., Zege, E.P., 2010b. Benchmark results in vector atmospheric radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer* 111, 1931–1946.
- Lebsock, M.D., L'Ecuyer, T.S., Stephens, G.L., 2007. Information content of near-infrared spaceborne multiangular polarization measurements for aerosol retrievals. *J. Geophys. Res.* 112, D14206. doi:10.1029/2007JD008535.
- Lee, K.H., Li, Z., Kim, Y.J., Kokhanovsky, A., 2009. Atmospheric aerosol monitoring from satellite observations: a history of three decades. In: Kim, Y.J., Platt, U., Gu, M.B., Iwahashi, H. (Eds.), *Atmospheric and Biological Environmental Monitoring*. Springer Science + Business Media B.V., pp. 13–38.
- Li, Z., Zhao, X., Kahn, R., Mishchenko, M., Remer, L., Lee, K.-H., Wang, M., Laszlo, I., Nakajima, T., Maring, H., 2009. Uncertainties in satellite remote sensing of aerosols and impact on monitoring its long-term trend: a review and perspective. *Ann. Geophys.* 27, 2755–2770.
- Lourakis, M.I.A., 2004. Levenberg–Marquardt nonlinear least squares algorithms in C/C++. <http://www.ics.forth.gr/~lourakis/levmar/>.
- Martonchik, J.V., Diner, D.J., Kahn, R., Verstraete, M.M., Pinty, B., Gordon, H.R., Ackerman, T.P., 1998. Techniques for the retrieval of aerosol properties over land and ocean using multiangle data. *IEEE Trans. Geosci. Remote Sens.* 36, 1212–1227.
- Mishchenko, M.I., Travis, L.D., Lacis, A.A., 2002. *Scattering, Absorption, and Emission of Light by Small Particles*. Cambridge University Press, Cambridge (UK).
- Nocedal, J., Wright, S.J., 2006. *Numerical Optimization*, XXII Springer Series in Operations Research and Financial Engineering 2nd edition. Springer-Verlag, NY. 664 pp.
- Rodgers, C.D., 2000. *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing Co., Singapore. 240 pp.
- Rozanov, A., Rozanov, V., Buchwitz, M., Kokhanovsky, A., Burrows, J.P., 2005. SCIATRAN 2.0 – a new radiative transfer model for geophysical applications in the 175–2400 nm spectral region. *Adv. Space Res.* 36, 1015–1019.
- Waquet, F., Cairns, B., Knobelspiesse, K., Chowdhary, J., Travis, L.D., Schmid, B., Mishchenko, M.I., 2009. Polarimetric remote sensing of aerosols over land. *J. Geophys. Res.* 114, D01206. doi:10.1029/2008JD010619.
- Zhai, P., Hu, Y., Trepte, C.R., Lucker, P.L., 2009. A vector radiative transfer model for coupled atmosphere and ocean systems based on successive order of scattering method. *Opt. Express* 17, 2057–2079.